DOCUMENT RESUME

ED 441 852 TM 030 887

AUTHOR Sarouphim, Ketty M.
TITLE Use of the DISCOVER Assessment for Identification Purposes: Concurrent Validity and Gender Issues.
PUB DATE 2000-04-00
NOTE 37p.; Paper presented at the Annual Meeting of the American Educational Research Association (New Orleans, LA, April 24-28, 2000).
PUB TYPE Reports - Evaluative (142) -- Speeches/Meeting Papers (150)
EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS American Indians; *Concurrent Validity; Elementary Education; *Elementary School Students; *Gifted; Hispanic American Students; *Identification; Mexican Americans; *Minority Groups; Multivariate Analysis; *Sex Differences
IDENTIFIERS *DISCOVER System; Raven Progressive Matrices

ABSTRACT

This study examined the DISCOVER (Discovering Intellectual Strengths and Capabilities through Observation while allowing for Varied Ethnic Responses) assessment (C. Maker, A. Nielson, and J. Rogers, 1994) as a concurrent measure of the Raven Progressive Matrices (J. Raven, J. Court, and J. Raven, 1977). It also investigated gender differences in mDISCOVER results. A secondary purpose was to determine the effectiveness of the DISCOVER assessment in reducing the problem of the underrepresentation of minority students in programs for the gifted. The sample consisted of 257 kindergarten, second, fourth, and fifth grade students, predominantly Navajo Indians and Mexican Americans. The results provide some evidence for concurrent validity and show that through use of the DISCOVER assessment 22.9%, of minority students were identified as gifted. A MANOVA, multivariate analysis of variance (gender by grade level), yielded no significant differences in the performance of males and females in all activities across grade levels. Chi-square tests revealed no overall significant gender differences between identification. The findings support the use of the DISCOVER assessment for identification purposes. (Contains 6 tables and 32 references.) (Author/SLD)

Reproductions supplied by EDRS are the best that can be made
from the original document.

ED 441 852

Use of the DISCOVER Assessment for Identification Purposes: Concurrent Validity and Gender Issues

Ketty M. Sarouphim

Lebanese American University

Paper presented at the Annual Meeting of the American Educational Research Association (April, 2000), New Orleans, Louisiana.

Dr. Ketty M. Sarouphim is Assistant Professor at the Lebanese American University, Byblos campus, Lebanon, Division of Social Sciences.

The author acknowledges the contribution of C. June Maker.

For correspondence with author, Please use this address: 475 Riverside Dr., Rm 1846, New York, NY 10115. E-mail: ksarufim@byblos.lau.edu.lb. Fax: 011961-9547-256

BEST COPY AVAILABLE

TM030887

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
☑ This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL HAS BEEN GRANTED BY

K. Sarouphim

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

Abstract

This study examined the DISCOVER assessment as a concurrent measure of the Raven Progressive Matrices. It also investigated gender differences. A secondary purpose was to determine the effectiveness of the DISCOVER assessment in reducing the problem of minority students' being under represented in programs for the gifted. The sample consisted of 257 kindergarten, second, fourth and fifth graders, predominantly Navajo Indians and Mexican-Americans. The results provided some evidence for concurrent validity and showed that through the use of the DISCOVER assessment 22.9% of minority students were identified as gifted. A MANOVA (gender by grade level) yielded no significant differences between the performance of males and females in all activities across grade levels. Chi-square tests revealed no overall significant gender differences in identification. The findings promote the use of the DISCOVER assessment for identification purposes.

Use of the DISCOVER Assessment for Identification Purposes: Concurrent Validity and Gender Issues

The issue of identifying gifted students from culturally diverse groups has received much attention in the literature (Baker, 1996; Maker, 1992; Clasen, Middleton, & Connell, 1994; Nielson, 1994; Scott, Perou, Hogan, & Gold, 1992). Several researchers have investigated why minority students are overrepresented in remedial programs and underrepresented in programs for the gifted (Clasen et al.; Gardner, 1992; Maker, 1993; Nielson, 1994). The often-cited causes for such practices are mostly traditional definitions of giftedness, narrow conceptions of intelligence, and the use of traditional assessment procedures for identification purposes, such as standardized IQ tests (Clasen et al.; Cummins, 1991; Maker, 1992; Samuda, 1991).

Much of the criticism has addressed the issue of fairness. Several studies on standardized tests have revealed gender, ethnic, and cultural bias (Baker, 1996; Johnson, 1994). Researchers and educators have identified four major sources of this bias: the norms used for test interpretation, inadequacy of formats, bias in content, and linguistically loaded items (Baker, 1996). Consequently, educators have called for the use of more adequate instruments for identification purposes, such as alternative assessment methods (Clasen et al., 1994; Cummins, 1991; Gardner, 1992; Maker, 1992).

Historically, giftedness has been associated with superior academic ability or achievement, measured by grade point average or IQ (Nevo, 1994). Terman's (1925) definition of gifted individuals as only those who scored in the top one percent in general intellectual ability on the Stanford-Binet Intelligence Test exemplifies how giftedness was viewed three quarters of a century ago. Evidence from recent publications indicates that the notion is being reconceptualized (Nevo, 1994). In 1972, a committee formed by the U.S. Office of Education (Marland, 1972) proposed a conception of giftedness which included not only abilities in the academic domain, but also in the performance domains. Children could be identified as gifted if they registered a high potential in the following areas: (a) general intellectual ability, (b) specific academic aptitude, (c) creative or productive thinking, (d) leadership ability, (e) visual and performing arts, and (f) psychomotor ability.

Renzulli's (1979) three-ring definition of giftedness is another reconceptualization of giftedness. He hypothesized that giftedness is an interaction between three clusters of basic traits: above-average general ability, high levels of creativity, and high levels of motivation (task commitment). Along the same lines, Maker (1993) postulated that creativity and intelligence are two components of the same construct. She contended that "creative problem-solving" is a characteristic of giftedness. According to Maker (1996), the key element in giftedness is the

ability to solve complex problems in the "most efficient, effective, or economical ways" (p. 44). Thus, in Maker's view, gifted individuals are both highly intelligent and creative; not only do they understand problems and discover solutions using the most efficient methods, they also find roblems and solve them creatively and effectively (Maker,1993, 1996).

In the same vein, the emergence of nontraditional theories of intelligence based on a broad conceptualization of intelligence has contributed to a reform of the concept as well. For example, Gardner (1983) defined intelligence as the multiple abilities that permit an individual to solve a problem or create a product that is valued within one or more cultural settings. In his book, *Frames of Mind*, Gardner (1983) rejected the unitary construct of intelligence and espoused a multidimensional definition in which he identified seven discrete intelligences: linguistic, logical-mathematical, spatial, interpersonal, intrapersonal, bodily-kinesthetic, and musical. More recently, Gardner (1999) has added one and a half intelligences to the previously identified seven; the eighth intelligence he labeled the "Naturalist" (botanist or sensitivity to the ecological environment) and the half intelligence he called the "Existentialist" (insight into the meanings of life and existence).

Performance-based assessments

The new conceptions of giftedness (e.g., Maker, 1993, 1996; Renzuli, 1979) and the reconceptualization of human intelligence (e.g.; Gardner, 1983, 1999)

have given rise to the development of performance-based assessments that have extended beyond the use of standardized tests (Clasen et al., 1994; Maker, 1996). Proponents of performance assessment see many benefits associated with this technique, such as testing students in lifelike situations, consideration of both process and product in evaluation, assessment of higher order skills, and use of appealing material (Frechtling, 1991). Specific to the assessment of culturally diverse groups, the advantages often-cited include (a) use of the dominant language of the person assessed; (b) coverage of broad and multiple areas such as those advocated by Gardner (1983) and Sternberg (1991); (c) performance based assessments do not yield scores that will be transformed into standard z-scores to be compared with the scores of the normative sample; rather, evaluation of the individual is based on the judgment of multiple observers or evaluators, such as independent observers, parents and peers; and (d) these methods are believed to be more fair and culturally bias-free in comparison with multiple-choice questions which might require knowledge and skills specific to the dominant culture (Baldwin, 1985; Maker, 1992).

The effectiveness of performance-based assessments has been investigated in several studies. For example, Clasen et al. (1994) conducted a well designed study in which they tested 433 minority and nonminority students using nontraditional multiple measures: problem solving, a free response drawing task, peer

identification, and teacher nomination. The results showed that 24% of the students tested were identified as gifted, and minority and nonminority gifted students were identified in proportion to their actual distribution in the schools. Peer and teacher nominations supported the art and problem-solving identifications. Also, the number of males and females identified corresponded closely to their proportions in the population. The researchers concluded that nontraditional measures may be more culture and gender fair than are traditional assessments.

In another study, Borland and Wright (1994), described an extensive method for the identification of economically disadvantaged students which included both qualitative and quantitative measures. Standardized tests as well as classroom observations, portfolio assessment, teacher nominations, and child interview were used for identification purposes. Validation data for two cohorts (K-2) yielded positive results. The researchers concluded that giftedness can be found in every school and that educators have no excuse for failing to identify gifted students from all backgrounds.

The DISCOVER assessment

Using the conceptual framework of Gardner's theory of multiple intelligences (1983) and Maker's definition of giftedness (1993), Maker, Nielson, and Rogers (1994) developed a performance-based assessment designed to identify gifted students among culturally diverse groups, called the DISCOVER assessment.

The acronym DISCOVER stands for Discovering Intellectual Strengths and Capabilities through Observation while allowing for Varied Ethnic Responses. For an extensive description of the DISCOVER assessment, see Sarouphim (1999).

The DISCOVER assessment is a relatively new instrument, consequently, only a few studies have examined its psychometric properties. Griffiths (1996) conducted two studies on the inter-observer reliability of the DISCOVER assessment. In the first study, two observers separately watched videotapes of five observation sessions of the Pablo® activity (spatial intelligence). Participants were 25 Navajo children ranging in age from 9 to 13 years. As they viewed the tapes, the researchers sketched the children's constructions and took notes in much the same way as the original observers in the tapes did. Then each of the researchers independently classified the children's problem-solving ability in Pablo® according to the four rating categories of Unknown, Maybe, Probably, and Definitely. Correlational analyses yielded positive and significant coefficients with the lowest being 0.69 ($p<0.05$) and the highest 0.81 ($p<0.01$), indicating a fairly high agreement among the three observers. Percentages of agreement ranged from 75% to 100%.

In the second study, participants were observed in a live setting. Six observers with different levels of experience (novices, moderate experience, and experts) watched the students perform three of the DISCOVER assessment activities (Pablo®,

Tangrams, and Storytelling) and recorded separate notes. Participants were 91 students ranging in age from 5 to 11 years old. Correlational analyses yielded positive and significant coefficients; the percentage of agreement between the researcher and all six observers ranged between 80 and 100% with the highest agreement being between the researcher and the expert observers and the lowest between the researcher and the novices. Also, the agreement among observers was 95 to 100% across all experience levels on the "Definitely" rating category. The researcher concluded that the DISCOVER assessment inter-observer reliability was high. Levels of observers' experience affect slightly, but not significantly their rating of students' problem-solving ability.

In another study, Seraphim (1997) investigated some aspects of the internal structure of the DISCOVER assessment checklist to assess construct validity. Participants were 368 American Indians and Mexican Americans taken from kindergarten, and fourth, fifth and sixth grades. Convergent and divergent validity of the checklist were assessed through correlations of observers' ratings of students' problem-solving ability in one activity and their rating of the same students in the other four activities. The results showed low and non significant inter-rating correlations, indicating that the checklist had high divergent validity. That is, students given high or low ratings in one activity were not necessarily given the same high or low rating

in the other activities, suggesting that each of the DISCOVER assessment activities measures a different intelligence. Analyses of gender differences revealed no significant differences in the numbers of males and females identified as gifted. The results indicated a good fit between the assessment and the theory of multiple intelligences, providing positive evidence for the construct validity of the DISCOVER assessment.

In a study with a purpose similar to the present investigation, Griffiths (1997) examined the comparative validity of the DISCOVER assessment with other measures. Thirty-four Mexican-American participants took the WISC-III, the Raven Progressive Matrices, and the DISCOVER assessment. Although overall ratings of students in the three assessments were strikingly different, analyses of separate activities corresponding to the different intelligences and students' profiles revealed high comparative validity indicating a close resemblance between the results of the DISCOVER assessment and the WISC-III and between the Raven's and the Pablo® activity of the DISCOVER assessment. Also, multiple regression analyses revealed that the DISCOVER assessment had higher predictive validity than either the Raven's or WISC-III, hence providing further evidence for the effective use of the DISCOVER assessment with minority students.

The primary purpose of the current study was to examine the validity of the DISCOVER assessment as a concurrent measure of

the Raven Progressive Matrices (Raven, Court, & Raven, 1977, 1988). Some investigators have suggested that the use of the Progressive Matrices with culturally diverse groups was appropriate (Jensen, 1980; MacAvoy, Orr, & Sidles, 1993) and leads to the identification of a higher proportion of minority children than traditional measures do (Mills & Tissot, 1995). Test-retest reliability for the Raven ranges between 0.71 and 0.92 and concurrent validity estimates are between 0.55 and 0.86 (Sattler, 1988). This inquiry also investigated gender differences in the use of the DISCOVER assessment. A secondary purpose was to determine whether users of the DISCOVER assessment would identify a larger pool of students than those using standardized tests and, thus, whether the use of the DISCOVER assessment would help reduce minority underrepresentation in programs for the gifted.

Method

Participants

The sample of this study consisted of 257 participants, predominantly from two minority groups: Navajo Indians and Mexican Americans. Participants were kindergartners, and second, fourth, and fifth graders taken from six schools located in the northern and southern parts of the state of Arizona. Most participants were from low socioeconomic groups as determined by their place of residence and participation in the free lunch program. Participants' grade, gender, and ethnicity distribution

is presented in Table 1.

Instruments

The instruments used in this study are the DISCOVER assessment and the Raven Progressive Matrices. The following is a brief description of each instrument:

The DISCOVER assessment. The DISCOVER assessment was designed to tap into individuals' problem-solving ability through five activities: Pablo® (spatial), Tangrams (spatial/logical-mathematical), Math (logical-mathematical), Storytelling (linguistic), and Storywriting (linguistic). The assessment consists of a series of tasks which students perform while being assessed by trained observers. To avoid observer bias, observers rotate at the completion of each activity so that each student is assessed only once (i.e., during one activity only) by the same observer. The following is a brief description of each activity:

Pablo®: The material for this activity consists of colored cardboard pieces of different shapes, designs, and sizes. Students are asked to make different constructions (e.g., animal, flowers, container) using the Pablo® pieces.

Tangrams: Each student is given a set of Chinese Tangrams (21 pieces of three different shapes: triangles of three different sizes, squares and parallelograms). Students are requested to make a geometrical shape (square in K-2 and triangle in grades 3-5) using as many Tangram pieces as possible; then each student is given a booklet of six puzzle sheets arranged in

ascending order of difficulty and asked to solve them.

Storytelling: Students are given an array of toys and are asked to either group the toys according to similarity in characteristics (K-2) or to describe one and then two of their toys using as many descriptors as possible (grades 3-8). Then students are asked to tell a story of their choice which incorporates some or all of the toys they have been given.

Storywriting: Students are asked to draw a picture which tells a story and verbally describe it (Kindergarten) or to write a story of their choice (grades 1-8).

Math: Worksheets consisting mostly of open-ended numerical problems are used to assess this intelligence (in kindergarten, Tangram pieces are used to assess the children's counting ability as well as their grasp of the concepts of "more" and "less").

Assessment procedures. Following the assessment, observers meet to discuss students' problem-solving abilities and classify their performance or strength in each of the activities according to a 4-category rating scale: Unknown, Maybe, Probably, and Definitely, with the last rating category being the highest and corresponding to superior problem-solving ability or giftedness. Usually, students given the "Definitely" rating category in at least two of the activities are identified as gifted; however, the identification criteria are flexible (e.g., in some schools, students given three "Definitely" ratings are identified as

gifted) and depend on the school district identification procedures and the nature and scope of programs for the gifted offered at each particular school.

Criteria for giftedness: To assign a rating, observers are guided by a checklist which they complete for each child. Items on the checklist represent superior problem-solving behaviors (process) and characteristics of products. For example, in Pablo®, observers note how the final construction was produced and whether the constructions are three dimensional, complex, and original, and incorporate many pieces. In Tangrams, observers note the number of puzzle sheets solved, the strategies used, the time it takes students to solve them and the number of Tangram pieces used to complete a square or a triangle. In Storytelling and Storywriting, observers look for fluency, plots, appropriate sequence of events, and the quality of words and sentences. In Math, strategies as well as the number of problems solved are taken into consideration.

Raven Progressive Matrices. Both the Raven Coloured Progressive Matrices (RCPM) and the Raven Standard Progressive Matrices (RSPM) are tests of nonverbal reasoning ability (Sattler, 1988). The RCPM, composed of 36 problems with colored matrices, is used with younger children, whereas the RSPM comprises 60 problems (divided into 5 sets of 12 items) with black and white matrices and is used with older children and adults. In both tests, the subject is required to find a missing piece which completes the

pattern in the displayed matrices.

Procedures

All participants took the DISCOVER assessment as well as the Raven Progressive Matrices (Raven et al, 1977, 1988). Kindergartners and second graders took the K-2 version of the DISCOVER assessment and the RCPM. Fourth and fifth graders took the grades 3-5 version of the DISCOVER assessment and the RSPM.

Results

Separate but identical analyses were performed on the checklists of students in each grade level. To determine concurrent validity, correlational analyses were performed between the participants' Raven scores and their DISCOVER ratings. R-squared was calculated to determine the percentage of shared variance between performance on the Raven and DISCOVER. For gender differences in activities across grade levels, a 2x4 MANOVA was conducted (gender by grade level). The ratings were coded as follows: 1 for "Unknown", 2 for "Maybe", 3 for "Probably", and 4 for "Definitely". Finally, chi-square tests of significance for gender by gifted participants (i.e., given the "Definitely" rating in at least two of the DISCOVER activities) were calculated to determine gender differences in identification.

Concurrent Validity

Correlations between the participants' Raven scores and their DISCOVER assessment ratings ranged between low and

nonsignificant, mostly for the Storytelling and Storywriting activities and moderate, high, and statistically significant for the other three activities (see Table 2). The lowest correlations were between participants' ratings in Storywriting and their Raven scores in all grade levels except in kindergarten and second grade, and the highest were between students' ratings in Pablo® and their Raven scores across grade levels except in kindergarten. A pattern of higher correlations for higher grade levels appeared, particularly in Pablo®.

Effect size as revealed by the variance explained in R-squared values yielded low to moderately high percentages, with the highest being 49% (R^2=0.49) between Pablo® and the Raven's in fifth grade and the lowest 0.86% (R^2=0.008) between Storywriting and the Raven's across the entire sample (see Table 3).

Gender Differences

By grade level. As shown in Table 5, the 2x4 MANOVA yielded non-significant *F*-tests indicating the absence of gender differences across grade levels in all activities of the DISCOVER assessment. Also, effect sizes were small as indicated by the low values of eta-squared. In general, the means of boys and girls ranged between a low and a high "Maybe", with few means reaching the "Probably" rating category (see Table 4). In kindergarten, fourth, and fifth grades, the performance of boys and girls was similar in all activities; in second grade, boys achieved higher ratings in all activities except Storywriting, but none of the

differences were significant.

In Pablo® and Math, boys achieved higher means across grade levels, but the differences were not significant. In the Tangrams, and Storywriting activities, the means were similar for both genders. In Storytelling, girls achieved higher means across grade levels, but the difference appeared non significant.

By gifted participants. As indicated in Table 6, 24.3% of kindergarten participants were identified as gifted, that is the boys and girls given the rating of "Definitely" in at least two of the DISCOVER assessment activities. A slightly lower percentage of students identified as gifted appeared in all other grade levels: second (23.4%), fourth (21.6%), and fifth (22.2%). A total of 22.9% of all participants was identified as gifted in the entire sample.

In terms of gender differences, no significant statistical differences were found between the number of boys and girls identified as gifted in all four subsamples (see Table 6) and subsequently across the entire sample, $X^2(1,257)=0.125$, ns.

Discussion

In this study, the purpose was to examine the validity of the DISCOVER assessment as a concurrent measure of the Raven Progressive Matrices. Another purpose was to investigate gender differences across activities and grade levels. A secondary purpose was to determine the effectiveness of the assessment in identifying higher percentages of minority students than

traditional standardized tests do. The results provided positive evidence for the concurrent validity of the DISCOVER assessment and showed that large percentages of participants were identified across the entire sample. Also, the 2x4 MANOVA on gender differences yielded non significant *F*-tests in all activities of the assessment across grade levels. Finally, no overall statistically significant differences were found in the numbers of boys and girls identified as gifted in each grade level and across the entire sample.

In this study, some evidence was revealed in support of the convergent and divergent validity of the DISCOVER assessment. The three activities of Pablo®, Tangrams, and Math require spatial and logical-mathematical reasoning; by the same token, both RCPM and RSPM are measures of nonverbal reasoning ability. Therefore, the significant correlations found between these three activities and the Progressive Matrices provide support for the concurrent validity of the DISCOVER assessment. Similarly, the low and nonsignificant correlations which appeared between the Storytelling and Storywriting activities and the Raven's Progressive Matrices provide the same kind of evidence (divergent validity) since RCPM and RSPM are not measures of verbal ability, whereas Storytelling and Storywriting were designed to assess linguistic intelligence. Evidence for convergent and divergent validity was accentuated by the R-squared values which yielded low percentages of shared variance between the activities of

Storytelling/Storywriting and the Raven's across grade levels and higher percentages of shared variance between the Pablo® activity and the Raven's in second, fourth, and fifth grades.

An interesting finding is the pattern of higher correlations for higher grade levels between the DISCOVER assessment and the Progressive Matrices. One explanation may be related to the different versions of the tests used. It appears that the problems proposed in the DISCOVER assessment for 3-5 grades and the RSPM are more similar than the K-2 version of the assessment and the RCPM. Further analyses are needed to confirm and clarify this finding.

A noteworthy finding is the absence of gender differences across grade levels and activities of the DISCOVER assessment. Moreover, no gender differences were found in the number of boys and girls identified as gifted across grade levels. Similar results were reported in other studies that investigated the effectiveness of performance-based assessments and in which no gender differences were found (Clasen et al., 1994; Plucker, Callahan, & Tomchin, 1996). The finding that girls did as well as boys on the overall tasks of the DISCOVER assessment may indicate that the instrument is mostly fair and does not discriminate against females or males.

In this study, a relatively high percentage of participants were identified as gifted. This finding is congruent with the results of other studies in which a performance-based assessment

was used as the instrument for identification. For example, in the study conducted by Clasen et al. (1994), the final pool of identified students included 24% of the participants. One possible explanation for the relatively large percentage of identified participants in the present study may be the grounded theory on which the DISCOVER assessment is based. Given the nature of multiple intelligences, the possibility of identifying gifted minority participants using the DISCOVER assessment is higher than that in traditional assessments in which a full scale IQ normed mostly on the majority population is used for identification procedures. Adherents of a full scale IQ claim that gifted individuals are those with extremely high scores (two or two and a half standard deviations above the mean), thus constituting three to five percent of the population. Hence, in their view, giftedness is unidimensional and of one kind only. However, if we embrace the view advanced in the theory of multiple intelligences, giftedness takes many forms and becomes multidimensional. Statistically, the probability of identifying gifted students through the use of the DISCOVER assessment is much higher than that found in traditional tests of intelligence. By definition, through the use of the DISCOVER assessment, an individual is identified as gifted if he or she is given the rating of "Definitely" in at least two of the activities. Given that the DISCOVER assessment is composed of five activities, each individual could be identified as gifted through ten different

combinations (i.e., Pablo® and Tangrams, Pablo® and Math, Pablo® and Storytelling, Pablo® and Storywriting, Tangrams and Math, Tangrams and Storytelling, Tangrams and Storywriting, Math and Storytelling, Math and Storywriting, Storytelling and Storywriting). Thus, the probability of identifying giftedness in the population is largely increased through the use of the DISCOVER assessment which might explain the high percentage of participants identified as gifted across grade levels in this study.

In this study, some evidence for the convergent and divergent validity of the DISCOVER assessment was revealed. However, compelling data supporting a strong statistical relationship between the DISCOVER assessment and the Raven's were not found. Why then would one use a complex instrument such as the DISCOVER assessment rather than a simpler one like the Raven's? Mainly for three reasons: First, because the multidimensional nature of the DISCOVER assessment enables the practitioner to assess a variety of intelligences, including linguistic ability measured both orally and in a written form. Secondly, because the appealing material and interesting tasks used in the DISCOVER assessment might motivate students to a better performance and reveal strengths that a paper-and-pencil test cannot reveal. Thirdly, because giftedness is not measured through percentile ranks, hence is not limited to the upper 3% of the students' population. However, one must always keep in mind

the purpose of assessing students and accordingly, use the test which best suits their interests. Indeed, providing students with the services that best meet their needs must remain the objective behind every assessment.

In sum, given the historically ineffective assessment of minorities and their underrepresentation in programs for the gifted, a change in assessment procedures is warranted. This study showed that the use of the DISCOVER assessment with culturally diverse groups may reduce the problem of minority underrepresentation in programs for gifted students. Also, evidence of the concurrent validity of the assessment provided support for its use. Moreover, the absence of gender differences may add the element of fairness to the DISCOVER assessment.

However, the limitations of this study must be kept in mind before drawing conclusions. One limitation is that the sample consisted of students from two culturally diverse groups only, Mexican-Americans and Navajo Indians; therefore, further research is needed with participants from other culturally diverse groups (e.g., Asians, African-Americans) to support these findings. Another limitation is that the participants belonged to lower grades; additional studies encompassing participants from upper grade levels are needed to support the use of the DISCOVER assessment with populations of different ages. Moreover, the concurrent validity of the linguistic activities of the DISCOVER assessment (Storytelling and Storywriting) needs to be examined

using measures of verbal ability with previously established validity. Finally, further studies on the reliability (e.g., test-retest, internal consistency) and construct validity of the DISCOVER assessment need to be conducted before sounding a call for the use of the assessment on a wider scale.

References

Baker, E. L. (1996). Introduction to theme issue in educational assessment. Journal of Educational Research, 89(4), 194-196.

Baldwin, A.Y. (1985). Programs for the gifted and talented: Issues concerning minority populations. In F. Horrowitz & M. O'Brien (Eds.), The Gifted and Talented: Developmental Perspectives (pp. 223-249). Washington, D.C.: American Psychological Association.

Borland, J. M., and Wright, L. (1994). Identifying young, potentially gifted, economically disadvantaged students. Gifted Child Quarterly, 38(4), 164-171.

Clasen, D. R., Middleton, J. A, & Connell, T. J. (1994). Assessing artistic and problem-solving performance in minority and nonminority students using a nontraditional multidimensional approach. Gifted Child Quarterly, 38(1), 27-37.

Cummins, J. (1991). Institutionalized racism and the assessment of minority children: A comparison of policies and programs in the United States and Canada. In R.J. Samuda, S.L. Kong, J. Cummins, J. Pascual-Leone, & J. Lewis. (Eds.), Assessment and placement of minority students (pp. 97-107). Kingston/Toronto: Intercultural Social Sciences Publications.

Frechtling, J. A. (1991). Performance assessment: Moonstruck or the real thing? Educational Measurement: Issues and Practices, 10(4), 23-25.

Gardner, H. (1999, July). Multiple intelligences for the new millennium Paper presented at the Eighth International Conference on Thinking, Edmonton, Canada.

Gardner, H. (1992). Assessment in context: The alternative to standardized testing. In B. R. Gifford, & M. C. O'Connor. (Eds.), Changing assessments: Alternative views of aptitude, achievement, and instruction (pp. 77-120). Boston: Kluwer.

Gardner, H. (1983). Frames of mind: The theory of multiple intelligences. New York: Basic Books.

Griffiths, S.E. (1997). The comparative validity of assessments based on different theories for the purpose of identifying gifted ethnic minority students. Unpublished doctoral dissertation. The University of Arizona, Tucson.

Griffiths, S. E. (1996). The inter-observer reliability of the DISCOVER problem-solving assessment. Unpublished manuscript, The University of Arizona, Tucson.

Johnson, N. E. (1994). Use of the WISC-R with disadvantaged gifted children: Current practice, limitations and ethical concerns. (ERIC Document Reproduction Services No. ED 368 097)

Macavoy, J., Orr, S., & Sidles, G. (1993). The Raven Matrices and Navajo children: Normative characteristics and culture fair application to issues of intelligence, giftedness, and academic proficiency. Journal of American Indian Education, Fall, 32-43.

Maker, C. J. (1996). Identification of gifted minority

students: A national problem, needed changes and a promising solution. Gifted Child Quarterly, 40, 41-50.

Maker, C. J. (1993). Creativity, intelligence, and problem solving: A definition and design for cross-cultural research and measurement related to giftedness. Gifted Education International, 9, 68-77.

Maker, C. J. (1992). Intelligence and creativity in multiple intelligences: Identification and development. Educating Able Learners, XVII(4), 12-19.

Maker, C. J., Nielson, A.B., & Rogers, J.A. (1994). Giftedness, diversity, and problem-solving. Teaching Exceptional Children, 27(1), 4-19.

Marland, S. P., Jr. (1972). Education of the gifted and talented (V.I.). Report to the Congress of the United States. Washington, DC: U.S. Government Printing Office.

Mills, C. J., & Tissot, S. L. (1995). Identifying academic potential in students from underrepresented populations: Is using the Ravens Progressive Matrices a good ides? Gifted Child Quarterly, 39(4), 209-217.

Nevo, B. (1994). Definitions, ideologies, and hypotheses in gifted education. Gifted Child Quarterly, 38, 184-186.

Nielson, A. B. (1994). Traditional identification: Elitist, racist, sexist? New evidence. Communicator: The Journal of the California Association for the Gifted, 24(3), 18-31.

Plucker, J. A., Callahan, C. M., & Tomchin, E. M. (1996).

Wherefore art thou, multiple intelligences? Alternative assessments for identifying talent in ethnically diverse and low income students. Gifted Child Quarterly, 40, 81-90.

Raven, J. C., Court, J. H., & Raven, J. (1988). Manual for the Standard Progressive Matrices. London: H. K. Lewis.

Raven, J. C., Court, J., H., & Raven, J. (1977). Coloured Progressive Matrices. London: H. K. Lewis.

Renzulli, J. S. (1979). What makes giftedness? Re-examining a definition. Phi Delta Kappan, 60, 180-184.

Samuda, R. J. (1991). Psychometric factors in the appraisal of intelligence. In R. J. Samuda, S. L. Kong, J. Cummins, J. Pascual-Leone, & J. Lewis. (Eds.), Assessment and placement of minority students (pp. 25-40). Kingston/Toronto: Intercultural Social Sciences Publications.

Sarouphim, K. M. (1999). DISCOVER assessment: A promising alternative for the identification of gifted minorities. Gifted Child Quarterly, 43(4), 244-251.

Seraphim, K. M. (1997). Observation of problem-solving in multiple intelligences: Internal structure of the DISCOVER assessment checklist. Unpublished doctoral dissertation, The University of Arizona, Tucson.

Sattler, J. M. (1988). Assessment of children. (3rd ed.). San Diego: Jerome M. Sattler, Publisher.

Scott, M. S., Perou, R., Urbano, R., Hogan, A., & Gold, S. (1992). The identification of giftedness: A comparison of white,

Hispanic and black families. Gifted Child Quarterly, 36, 131-139.

Sternberg, R. J. (1991). Giftedness according to the triarchic theory of human intelligence. In N. Colangelo, & G.A. Davis (Eds.), Handbook of Gifted Education (pp. 45-54). Boston: Allyn & Bacon.

Terman, L. (1925). Mental and physical traits of a thousand gifted children. In L. Terman (Ed.). Genetic studies of genius (Vol 1). Stanford: Stanford University Press.

Table 1

Participants' Grade, Gender and Age Distribution

	Kindergarten	Second	Forth	fifth	Total
			Gender		
Male	39	25	16	36	116
Female	35	22	30	54	141
Total	74	47	46	90	257
			Ethnicity		
Navajo	42	2	15	55	114
Hispanic	28	39	22	30	119
Anglo	4	6	9	5	24
Total	74	47	46	90	257

Table 2

Correlations Between Participants' Raven Scores and their DISCOVER Ratings

	Kindergarten (n=74)	Second (n=47)	Fourth (n=46)	Fifth (n=90)	All (n=257)
Pablo®	0.251*	0.506**	0.613**	0.704**	0.579**
Tangrams	0.351*	0.398**	0.495**	0.395**	0.409**
Math	0.264*	0.311*	0.376*	0.357**	0.311**
Story	0.297*	0.120	0.294	0.206	0.108
Writing	0.334*	0.276	0.139	0.198	0.093

*$p < 0.05$. **$p < 0.01$.

Table 3

R-Squared for Correlations Between the DISCOVER Ratings and The Raven's Scores

	Kindergarten	Second	Fourth	Fifth	All
Pablo®	0.063	0.256	0.375	0.495	0.335
Tangrams	0.123	0.158	0.245	0.156	0.16
Math	0.069	0.096	0.141	0.127	0.096
Story	0.088	0.014	0.086	0.042	0.011
Writing	0.111	0.076	0.019	0.039	0.008

Table 4

Mean ratings of boys and girls in Each DISCOVER Activity Across Grade levels

Activity	Mean		SD	
	Boys	Girls	Boys	Girls
		Kindergarten		
Pablo®	2.82	2.65	0.94	0.87
Tangrams	2.12	2.20	0.73	0.83
Math	2.76	2.57	0.90	1.03
Story	2.15	2.28	1.08	1.07
Writing	2.80	2.57	0.80	0.70
		Second		
Pablo®	2.92	2.81	0.70	0.66
Tangrams	3.00	2.63	0.76	0.84
Math	3.04	2.90	0.67	0.75
Story	2.72	2.71	0.84	0.71
Writing	2.64	2.90	0.81	0.81

Table 4 (continued).

Activity	Mean		SD	
	Boys	Girls	Boys	Girls
		Fourth		
Pablo®	3.37	2.63	0.50	0.69
Tangrams	2.50	2.60	0.81	0.81
Math	2.83	2.69	0.93	0.70
Story	2.75	2.92	1.06	0.91
Writing	2.60	2.58	1.18	0.88
		Fifth		
Pablo®	2.87	2.86	0.68	0.72
Tangrams	2.97	3.09	0.87	1.01
Math	2.87	2.61	0.87	1.03
Story	2.31	2.83	0.79	0.89
Writing	2.55	2.85	0.91	0.92

Table 5

Multivariate Analysis of Variance and Effect Size for Gender by Grade Level

	F	*P*	Eta^2
	Kindergarten		
Pablo®	*F*(1,57)=0.498	0.483	0.008
Tangrams	*F*(1,57)=0.002	0.964	0.000
Math	*F*(1,57)=1.323	0.254	0.021
Story	*F*(1,57)=0.010	0.922	0.000
Writing	*F*(1,57)=1.383	0.244	0.022
	Second		
Pablo®	*F*(1,40)=0.291	0.592	0.007
Tangrams	*F*(1,40)=2.517	0.120	0.054
Math	*F*(1,40)=0.403	0.529	0.009
Story	*F*(1,40)=0.001	0.981	0.000
Writing	*F*(1,40)=0.836	0.366	0.019

Table 5 (continued).

	F	*P*	Eta^2
		Fourth	
Pablo®	*F*(1,29)=2.759	0.122	0.149
Tangrams	*F*(1,29)=0.008	0.928	0.000
Math	*F*(1,29)=0.240	0.627	0.007
Story	*F*(1,29)=0.060	0.808	0.002
Writing	*F*(1,29)=0.214	0.647	0.006
		Fifth	
Pablo®	*F*(1,68)=0.005	0.942	0.000
Tangrams	*F*(1,68)=0.077	0.782	0.001
Math	*F*(1,68)=2.316	0.123	0.057
Story	*F*(1,68)=2.506	0.111	0.034
Writing	*F*(1,68)=1.441	0.234	0.020

Table 6

Chi-square Tests of Significance for Gender by Gifted Participants Across Grade Levels and for the Entire Sample

Grade	Boys		Girls		All		df	X^2
	n	%	n	%	n	%		
Kindergarten	10	17.9	8	22.8	18	24.3	1	0.07
Second	8	32.0	3	13.6	11	23.4	1	2.20
Fourth	5	31.2	5	16.6	10	21.6	1	1.30
Fifth	9	25.0	11	20.3	21	22.2	1	1.09
All	32	27.5	27	19.1	59	22.9	1	1.89